

Social values and institutional change: An experimental study

KLARITA GËRKHANI*

European University Institute, Florence, Italy

JACQUELINE VAN BREEMEN**

University of Amsterdam, Amsterdam, The Netherlands

Abstract: We experimentally investigate whether and how individuals change formal institutions governing an organization. The focus is on formal rules prescribing individual cooperation to achieve a collective goal. Our analysis accounts for the role of social norms and individuals' social values. We observe that prosocial individuals –who value cooperation and have a conflict between this value and existing rules allowing for low cooperation– attempt to change this rule. In line with our theoretical discussion, we also find that prosocial individuals first try to change the institutional environment by changing social norms. If this fails, these individuals change formal rules directly.

* Klarita.Gerxhani@eui.eu

** Jacquelinevb29@hotmail.com

1. Introduction

An organization is an environment where institutions play an important role in structuring members' behavior.¹ In this article, we combine notions from economics, sociology and psychology to investigate institutional change within organizations. We consider both formal institutions (like formal rules) and informal institutions (like social norms)^{2,3} and employ Ostrom's (1990: 140) definition of institutional change as "A change in any rule affecting the set of participants, the set of strategies available to participants, the control they have over outcomes, the information they have, or the payoffs". In particular, we study whether and how individuals change the formal rules of their organization. We focus specifically on the roles of pre-existing formal rules, social norms, and social values in the process of this change.

In line with both the bounded rationality perspective and analytical sociology, our conceptual framework is based on the assumption that individuals act in ways to find optimal outcomes while being constrained or enabled by rules, norms and values. In other words, our starting point is that (institutional) structure and agency are not opposing forces but, instead,

¹ We adopt North's (1990: 4–5) view on organizations as "groups consisting of individuals achieving objectives bounded by a common purpose". Like Hodgson (2006: 2), we see institutions as "systems of established and prevalent social rules that structure social interactions".

² We distinguish between 'formal institutions' and 'informal institutions' for analytical reasons. Formal institutions are typically defined as rules or regulatory policies that are explicitly written down and enforced, either by third-party sanctions or by the organization itself. Informal institutions are social norms, prescribing the expected appropriate behavior in a specific situation, and they are informally enforced through (social) rewards or punishment by peers (Elster 1989). These are institutions type 3 and 4 in Voigt's (2018) recent article on informal institutions (Table 1: 3).

³ Henceforth, we use 'formal rules' and 'social norms' when referring to formal institutions and informal institutions within organizations, respectively.

interdependent and mutually constitutive (e.g. Giddens, 1984), and that “the interplay of both is required to understand how institutions are formed and sustained” (Hodgson, 2006: 19).

The environment we consider is one where a team of individuals needs to cooperate in order to achieve a collective outcome that benefits them all.⁴ The benefits of the collective outcome are available to everyone in the team, independent of the individual cooperation decision.

Hence the team faces a social dilemma.

Our framework combines three levels of analysis. First, we investigate the effects of institutional structure on agency (the ‘macro-to-micro link’: Coleman, 1990) by focusing on how formal rules affect behavioral outcomes. The formal rules we consider are ‘cooperation rules’ which prescribe individual cooperation to achieve a team’s collective goal. In this way, organizational rules may guide and structure the interaction within teams (Sewell, 1998). On the one hand, such rules might prescribe cooperative behavior towards a common goal. On the other hand, the rules might be lenient and allow individual team members to free ride on the cooperative behavior by others, leading to a breakdown of cooperation (Olson, 1965). To account for both possibilities, our framework distinguishes between two organizational environments that vary with respect to the extent of mandatory cooperation (i.e., the extent to which an existing rule leaves the decision to act cooperatively at the discretion of the individual).

Second, we examine agency by focusing on interactions at the individual and team levels (the ‘micro-to-micro link’: Coleman, 1990). We assign an important role to individual social values in exploring these micro-level dynamics. These values have been argued to be important (e.g. Parsons, 1966) and ‘socially meaningful’, because they connect individuals with social structures (Hitlin and Pinkston, 2013: 320).

⁴ Of course, many organizations are more complex than teams. We choose the simplest form of an organization, in order to provide a benchmark for further investigation under more complex organizational structures.

Finally, we investigate the role of agency on institutional structure (the ‘micro-to-macro link’: Coleman, 1990) by looking at the effect of micro-level interactions on institutional changes from an existing formal cooperation rule to a new one.

The research centers on a laboratory experiment. The experimental approach enables us to disentangle formal rules, social norms, social values, and the underlying mechanisms of institutional change (Jackson and Cox, 2013). In particular, the experimental approach provides us with tools to eliminate confounding variables that hinder causal inferences in natural settings.

Our study relates to the existing experimental literature on sanctioning and rewarding systems in social dilemmas. Like ours, this literature addresses both first- and second-order social dilemmas (Ostrom, 1990). It argues that when a group faces a social dilemma there is a potential role for an institution to promote cooperation. The success of the institution often depends on individual efforts, however, like the willingness to punish free riders. This is what may create a second-order dilemma, where individuals can free ride in the hope that others will make the effort required for the institution to work. The main focus of this literature is on enforcement mechanisms, such as formal and informal systems of sanctioning and rewarding, which can be endogenously chosen or exogenously imposed (e.g., Fehr and Gächter, 2000; Reuben and Riedl, 2013; Sefton *et al.*, 2007; Yamagishi, 1986).

Our contribution to this literature is threefold. First, we address a different second-order collective good dilemma. Instead of studying enforcement mechanisms, we focus on formal institutions themselves and the likelihood of them being changed in order to solve the first-order collective good dilemma.⁵ Doing so recognizes an important aspect, which is that individuals interacting in social dilemmas do not operate in a vacuum; their actions are structured by pre-existing formal rules. Second, we try to unravel the social mechanisms

⁵ See Janssen *et al.* (2008) for a similar approach.

underlying this institutional change by focusing on micro-level interactions, where individual social values take a prominent role. Our final contribution relates to the institutional literature (e.g., Brousseau *et al.*, 2011; Greif and Kingston, 2011; Kingston and Caballero, 2009). This literature distinguishes between on the one hand institutions as rules that are kept distinct from their enforcement, and on the other hand institutions as equilibria of the interaction amongst individuals. By investigating why individuals follow or change rules, we bring together these ‘institutions-as-rules’ and the ‘institutions-as-equilibria’ perspectives as advocated by Greif and Kingston (2011: 15).

2. Institutional change

Structural theories of path dependence have typically predicted institutions to be resistant to change, in particular due to the ‘entrapment of actors’ (Pierson, 2000: 253). Agency theories, on the other hand, have argued that actors can endogenously change institutions. This is the starting point of our study. Note, however, that pursuing institutional change may be costly (Yamagishi, 1986). Think for instance of organizational sanctions such as a reprimand or a fine (Morrison, 2006), or informal sanctions such as ostracism or derogative reactions by other members of one’s team (Nee and Ingram, 1998). It is then not a priori obvious that change will occur. This is the situation we are interested in. For this reason, we make the initiation of institutional change costly. A first important question, is *who*, if anyone, is willing to bear these costs? To study this, we distinguish between two scenarios. In the first, all individuals agree that the rule change is desirable. In the second scenario, only some individuals feel this way.

First, assume that desirability of the rule change is shared by all. That means for any individual, that if she were the only one who could initiate change, she would do so. When more individuals can initiate change, however, the costs involved imply that each individual

prefers that someone else takes the initiative. This then constitutes a volunteer's dilemma (Diekmann, 1985). All agree that an act is desirable, but everyone would like somebody else to bear the costs (i.e., initiating institutional change). The question, then, is who will volunteer?⁶ Previous studies have considered, i.a., how volunteering depends on the (a)symmetry of costs and benefits of the act or on how cost sharing affects the number of volunteers (Chen *et al.*, 2013; Weesie and Franzen, 1998). We look beyond material motives and argue that an individual's inclination to volunteer for a rule change depends on her social value orientation and its relationship to the pre-existing rule she wants to change.

This takes us to the second scenario, where there is no agreement about the desirability of a rule change. Some prefer to keep everything as is, while others want to change the rule. In this case, the decision to attempt the rule change no longer constitutes a volunteer's dilemma. We will argue that here, too, social values play an important role in explaining who will initiate change.

In what follows, we therefore first introduce social values and then discuss for both scenarios how they predict who will initiate a change.

Social values

Individual social values (a.k.a. internalized moral norms: Greif and Kingston, 2011) are typically defined as people's generalized beliefs regarding the desirability of conducts or end-states (e.g., Hitlin and Pinkston, 2013). They are formed through socialization within families, friendships, social class, or working environments (Parsons, 1966). It is important to note that values and preferences are distinct concepts. While individuals' value orientations

⁶ For simplicity, we focus on the situation where a single individual can initiate change. This suffices to describe the dilemma involved. In our experiment any single individual can indeed initiate change, though whether she succeeds depends on decisions by all.

capture their general principles that can be applied to a range of situations, their preferences may differ depending on particular outcomes, specific actions, or context (e.g., Tao and Au, 2014). Values may affect one's preferences or attitudes but are more general than that. Relevant examples of values are self-transcendence ('prosocial') values, which reflect concern for others' welfare and motivate cooperative behavior (Van Lange *et al.*, 1997). When applied to specific environments, these values may lead to social or other-regarding preferences as these are typically understood in the experimental economics literature (see Brousseau *et al.*, 2011 for an overview). Self-enhancement ('proself') values, on the other hand, relate to pursuing self-interest and motivate the accumulation of personal wealth and power. Values have been shown to be vital in guiding evaluation of alternatives and shaping behavioral choices (see for an overview: Tao and Au, 2014).

Studies in social psychology argue that when an individual faces social values and formal rules that are incongruent, she may experience a conflict between the personal and social identity, making the shared social identity less salient (Turner *et al.*, 1994). This, in turn, may diminish support for the formal rules that embody the social identity. We conjecture that this situation increases the likelihood that such individuals will attempt to change the formal rules they are subjected to.

First, we consider the volunteer's dilemma that occurs when everyone agrees that an existing rule should change (but, rationally speaking, no one wants to volunteer to initiate this change). Applying the general reasoning from social psychology to this scenario implies that individuals experiencing a conflict between their values and an existing rule are more likely to volunteer to bear the costs of initiating a change in this rule. If, for example, an existing rule leaves the decision to act cooperatively at the discretion of the individual, then those with a proself value have no conflict and might refrain from such actions as long as the rule is in place. Prosocials, however, will experience a conflict between this rule and their intrinsic

value to act cooperatively. We can then expect that it will be the prosocials and not the proselfs who will attempt a rule change. More generally, we propose that those with a higher conflict between the existing rule and their value are more likely to volunteer in a volunteer's dilemma. Though the prediction that prosocials are more likely to initiate change in an uncooperative rule seems intuitive, to the best of our knowledge no previous study has addressed the question of how volunteering in a volunteer's dilemma correlates with one's value orientations.⁷

For the second scenario, where only some individuals would like to see the rule changed, we again consider the case where an existing rule leaves the decision to act cooperatively at the discretion of the individual. Proselfs will typically have no desire to change the rule, because it enables them to pursue their self-interest with no restrictions.⁸ The prosocials, on the other hand, are more concerned about the enhancement of joint outcomes, which are less likely to be realized under a less mandatory cooperation rule. Such a rule is therefore in conflict with their values and this conflict will make the prosocials more likely than the proselfs to initiate a change. Hence, also for this scenario, it is those with a conflict between the existing rule and their value orientation that are most likely to initiate a rule change.

This reasoning leads to our first theoretical prediction (TP):

⁷ There are many studies showing that prosocials cooperate more than proselfs in social dilemma's (e.g. Offerman *et al.*, 1996; Van Lange *et al.*, 1997). These results, however, do not directly carry over to the strategic environment of the volunteer's dilemma. To see this, consider a group of four players, three of which are forced to choose the selfish option (contribute nothing in a public goods game or do not volunteer in the volunteer's dilemma). If the fourth player is selfish, she will contribute nothing in the public goods game, but she will volunteer in the volunteer's dilemma. This illustrates how the private incentives differ between the two.

⁸ After introducing our experimental design, we will derive specific conditions on when proselfs do or do not prefer to change the existing rule (cf. fn 16).

Theoretical Prediction 1. Individuals, whose social values are in conflict with the formal rules, are the ones most likely to attempt to change these rules.

TP1 predicts *who* will initiate institutional change. At first sight, one might think that those whose values are in harmony with the existing rules have no reason to change the latter. Recall from our discussion above, however, that values can be different from preferences. Those who experience rule-value harmony may still have a preference for a different environment because this might make them better off in a specific context. In our experiment, values indeed capture generalized principles. For example, a proself value describes a general tendency to choose the best for oneself, irrespective of what others do. This is in harmony with a rule that allows for low contributions. Still, there are cases where a proself might initiate a change to a more mandatory cooperation rule. This could occur, for example, if she expects that very few people will contribute under an uncooperative rule, hence she anticipates to earn more under a more cooperative rule, even if this requires her to give up the option of contributing little. TP1 implies, however, that such rule-change attempts by those with rule-value harmony are less likely than attempts by individuals with a conflict.

We now proceed to discuss *how* they do so. For this, the social norms governing behavior in an organization play a central role.

Social norms

An organization's social norms and formal rules form the structure that specifies not only how individuals in the organization ought to interact to receive rewards like status or salary, but also how they "compete" for the reproduction of those institutions (Rand and Nowak, 2013: 413). Thus, to understand how individuals with conflicting values attempt to change existing formal (cooperation) rules, we need to also take into account how the organization's social norms interact with its formal rules.

We thus focus specifically on the individuals experiencing a conflict between their individual values and the organization's formal rules. These are the individuals that we have predicted to attempt to change the formal rule (TP1). Following Nee and Ingram (1998: 35), we consider two possibilities with respect to the prevailing social norm. Either this norm aligns with the individual values or it is in agreement with the formal rule. In the former case, Nee and Ingram argue that individuals will typically find a way to circumvent the formal rules via "a decoupling of the informal norms and the formal rules of the organization". Thus, both are maintained, but formal rules exist only *de jure* and social norms *de facto*. In practice, no formal institutional change is needed because individuals informally agree that one ought to behave differently than formally prescribed.

The second possibility is that existing social norms are in agreement with formal rules but in conflict with individual values. In this case, individual interactions within organizations may give rise to 'opposition norms', which are resistant to the formal rules (Nee and Ingram, 1998). Opposition norms might undermine formal rules and this leads to institutional change. Thus, an individual facing a rule-value conflict, where the prevailing norm sides with the rule, has two options. She can initiate a change in the rule itself or she can try to make the rule irrelevant by changing the social norm into an opposition norm. The former will face more resistance than the latter, precisely because one is not only addressing an adverse rule, but also an adverse existing norm. We therefore predict that individuals in this situation will first try to change the norm into an opposition norm.

The attempted change to an opposition norm might fail, however. In this case, individuals with a conflict may attempt to change the formal rule directly if they "have substantial autonomy" (Ostrom, 1990: 21). This is most likely in organizations characterized by a democratic structure where all individuals have the ability and the right to participate in rule making (Jacob 2015). In such a democratic environment, an individual can directly

challenge an existing rule via voting (Walker *et al.*, 2000). When binding for all involved, voting on rule change may provide a successful way to establish socially optimal outcomes (Hauser *et al.*, 2014: 220). TP2 summarizes this discussion.

Theoretical Prediction 2.

- a) Individuals who have a value-rule conflict, in an environment where the social norm aligns with the formal rule, first address this conflict by attempting to change the corresponding social norm.
- b) If the social norm does not change, they will attempt to directly change the rules by voting.

Communication

Finally, we consider communication as a characteristic of the environment that may affect the chances of success after individuals initiate a rule change. Communication plays an important role in achieving common goals. Experimental research shows that it improves the ability of groups to overcome issues of cooperation (for an overview, see: Ostrom, 2005). Moreover, one of the prerequisites for institutional change is the presence of a minimum coalition in favor of change (DiMaggio, 1988; Ostrom, 1990). The formation of a coalition is facilitated by the ability to communicate not only with the like-minded but also with opponents (to persuade them). We therefore propose that any attempt to change a rule is more likely to succeed if individuals can contest it through debate, which can be facilitated by communication. This leads to the following prediction.

Theoretical Prediction 3. A rule change is more likely when individuals are able to communicate than without communication.

3. Experimental procedures and design

The experiment was conducted in June and October 2014 at the CREED Laboratory of the University of Amsterdam and consisted of a pre-study (explained below) and a ‘main experiment’. 220 individuals participated in the main experiment.⁹ We have data from a total of 215 participants (124 men and 91 women, $M_{\text{age}} = 22.62$, $SD = 3.16$). They participated in one of 17 sessions of the main experiment, in which they were randomly assigned to one of our four treatments cells. Sessions lasted approximately 90 minutes and participants earned on average €29.30 including a €7 show-up fee.

Procedures and tasks

Upon arrival, participants were randomly seated at separated computer cubicles, and were asked to complete multiple tasks, before participating in two public goods games (PGG1 and PGG2). They were informed that the experimental session consisted of multiple tasks and rounds, but they did not know how many rounds or which tasks would follow. Instructions were given at the beginning of each task.¹⁰ All participants received the same general instructions, aside from treatment-specific information.

The first task measures participants’ *risk attitudes* by asking “How willing are you to take risks, in general?”, with answers on a 11-point Likert scale ranging from 0 (not at all) to 10 (very much).¹¹ We measure risk attitudes because an individual’s willingness to take risks may positively influence the likelihood of organizational rule breaking (Morrison, 2006). The second task assesses participants’ *social value orientations* using the triple-dominance

⁹ Due to a computer crash in the 12th session, the data for five participants were lost.

¹⁰ See online Appendix A for a translated version

(<https://www.eui.eu/Documents/DepartmentsCentres/SPS/Profiles/Gerxhani/Appendix-AandB.pdf>).

¹¹ Dohmen *et al.* (2011) report high correlation between survey responses of this type and incentivized risk tasks like Holt and Laury (2002).

measure (TDM; Van Lange *et al.*, 1997). The TDM consists of nine items, each containing three distinct outcome distributions assigning points to oneself and to an (anonymous) paired other.¹² For example, a participant is asked to make a choice between (own earnings; other's earnings) options A=(480;80), B=(540;280), and C=(480;480). Depending on their choices, the TDM classifies participants into one of three categories – competitive (in this case, option A), self-interested (B), or prosocial (C). To isolate the rule-value conflict on the one hand, from 'internal' conflicts between an individual's values, on the other, we focus on participants that show 'consistent' value orientations, i.e. those that make the same kind of choice nine out of nine times. We therefore categorize a participant as *consistently* competitive, proself or prosocial if all of the choices can be attributed to the motive concerned. We measure the participants' social value orientations twice, before and immediately after PGG1, which we refer to as SVO1 and SVO2.

The third task measures the *social norm*, that is a general convention (Miller, 2008), of what one "ought to do" in a PGG. We do so by employing an adapted version of the method introduced by Krupka and Weber (2013).¹³ To avoid the measurement of a social norm interacting with participants' decisions in the main experiment, we conducted a pre-study in which participants were presented with the instructions of the PGG to be used in the main experiment.¹⁴ We asked the participants in the pre-study to predict which contribution to the team project would be considered most 'socially appropriate' by indicating a number between 0 and 10 points. Any participant who chose the modal response earned 500 points. By rewarding an estimation of the modal response, we did not elicit the participants' own

¹² We use anonymous RING matching, such that for participant 1 "the other" is participant 2, for 2 it is 3, and so forth. For the last participant "the other" is participant 1.

¹³ See Voigt (2018) for a call to rely more on experiments in measuring informal institutions like social norms.

¹⁴ The pre-study was executed in May 2014 at CREED with 23 participants. The session lasted approximately 20 minutes. Participants earned €8.52 on average, including a €7 show-up fee.

preferences but rather asked them to match the responses of others (i.e., we asked them to indicate what they thought was a ‘shared understanding’ of what one ought to contribute to the PGG). In this way, the participants play a coordination game in which they have to anticipate the extent to which others rate an action as most socially appropriate (Krupka and Weber, 2013). Our norm measure captures two important features of a social norm. The first concerns what is considered socially appropriate and the second addresses the social consensus (Bicchieri, 2006; Krupka and Weber, 2013).

The pre-study social norm was that one ought to contribute five points (out of 10). For the norm measurement in the main experiment, we present participants with the same pre-study instructions and ask them to indicate their own beliefs about which contribution to the team project the pre-study participants had regarded as most socially appropriate. Again, we provide a 500-point bonus for correctly estimating the pre-study’s modal response. This task serves to measure the beliefs held by participants in the main experiment about the prevalent norm. Here again, the modal response was five points. Subsequently, we disclose to all participants in the main experiment that a social norm of contributing five points to the team project had been considered most socially appropriate by those who took part in the pre-study. This revelation of the modal response serves to create a shared understanding of the prevalent norm for the environment of the main experiment, which captures the third important feature of a social norm (Fehr and Gächter, 2000).

Public goods game

The core of our experiment consists of two public goods game –PGG1 and PGG2– of 10 rounds each. After finishing the tasks described above, the participants are randomly assigned to a team of five for PGG1 (anonymous matching), and remain in this team throughout all rounds as well as in PGG2. Participants are unaware during PGG1 that a second public good

game will follow. In each round of every game, each individual is endowed with a fixed 10-points resource, which they can (partially or completely) contribute to a “team project”. The term “team project” was purposefully chosen to mimic a simple organizational environment. In each round the total team contribution is announced to the team members, but individual contributions remain unknown to others. Participants are informed that each point contributed to the team project yields a marginal per capital return (MPCR) of 0.4 points to every team member, irrespective of their own contributions. This MPCR of 0.4 yields individual earnings per round of:

$$10 \text{ points} - \# \text{ points contributed to the team project} + 0.4 * \text{total team contribution}.$$

Treatments

We use a full-factorial 2 x 2 between-subject design. In the first treatment –*Contribution rule*– we vary the formal rule prescribing a minimal mandatory contribution level in PGG1. The second treatment –*Communication*– varies whether or not communication with other group members is allowed.

In PGG1 the *Contribution rule* is either *Rule 2*, where a minimum contribution of 2 points is imposed in each round; or *Rule 8*, where the minimum contribution is 8 points. These formal rules are strictly enforced: participants are unable to contribute fewer points than the treatment calls for.¹⁵ Note that the *Rule 8* environment leaves the participants with very little freedom to choose (they can contribute 8, 9, or 10 points per round). As clarified below, this provides us with a benchmark to which we can compare the *Rule 2* environment.

After participating in PGG1, we first measure SVO2. Subsequently, all participants are given the opportunity to attempt a change of the existing contribution rule (*Rule 2* to *Rule*

¹⁵ Our rationale for doing so is that we are not interested in the role of enforcement mechanisms on institutional change. We therefore keep this aspect fixed throughout the experiment.

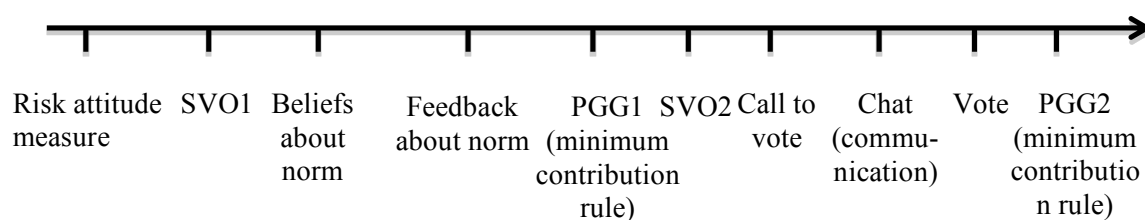
8, or vice versa, depending on the treatment) by initiating a call to vote. This is implemented by asking the participants to indicate “yes” or “no” to the question of whether they would like to call for a vote to change the existing contribution rule. In line with our theoretical discussion, a single individual ‘volunteering’ to initiate the vote suffices. The initiation of a vote costs 10 points, irrespective of others’ decisions. This amount reflects the observation that challenging an existing rule in the world outside of the laboratory is costly. If no team member calls for a vote, no voting procedure takes place and the minimum contribution level of PGG1 stays in place. If at least one team member initiates a call to vote –which is our measure of *rule-change attempt*– then a voting procedure starts, where the minimum contribution level of PGG2 is decided by majority rule (three out of five votes in favor of the change are required). Our measure of an *actual rule change* is thus defined as the switch from the rule implemented in the PGG1 to the opposing contribution rule.

In the treatments with communication, team members are allowed to communicate for 90 seconds through a chat box before the actual vote takes place. Participants are informed about this chat opportunity before they are given the possibility to call for a vote. In the treatments without communication, participants do not receive any additional information and vote without the possibility to communicate beforehand.

After a call to vote and the voting itself, PGG2 starts, with the same contribution rule as before if there was no call to vote or if the change option was rejected by a majority, and with a new contribution rule if there was a call for vote and if the majority supported the change.

Each session concludes with a short questionnaire on the participants’ socio-demographic characteristics. The experimenter then calls each participant separately to the back of the room to anonymously receive their payment. Figure 1 summarizes the timeline of the experiment.

Figure 1. Timeline of the Experiment.



Payoffs

Participants are paid in cash privately at the end of each session. Earnings are in “points” and exchanged for euros at a rate of €1 per 100 points. Earned points consist of: a possible bonus for the norm measurement; one randomly determined choice for each of the SVO measures; points earned in PGG1 and PGG2; and (negative) points if they initiated a call to vote.

Testable hypotheses

The theoretical predictions presented in the previous section straightforwardly yield hypotheses (H) that can be tested with the data from our experiment.¹⁶ As argued above, we

¹⁶ Recall that we distinguished between two scenarios, depending on whether or not all participants agree that change is desirable. Here we show that proselves may or may not want to change from *Rule 2* to *Rule 8* in our experiment. Proselfs will want to change if they expect to obtain a higher payoff under *Rule 8*. Using the parameters of our experiment, assume that a proself expects the four co-members to contribute on average x under *Rule 2* and on average 8 under *Rule 8*. She herself will contribute nothing under *Rule 2* and 8 under *Rule 8*. Her expected earnings under *Rule 2* are then $10 + 0.4 \cdot 4 \cdot x$. Under *Rule 8*, she expects to earn $2 + 0.4 \cdot 5 \cdot 8 = 18$. She will prefer not to change the rule whenever $10 + 1.6 \cdot x \geq 18$, or $x \geq 5$. This means that the call to vote constitutes a volunteer’s dilemma whenever the proselves expect the others to contribute on average less than 5. This also means that no volunteer’s dilemma is involved if the proselves expect others to adhere to the norm of contributing 5.

categorize those participants with *consistently* proself or prosocial choices as experiencing a conflict.

H1 (based on TP1).

In *Rule 2*, consistent prosocials are more likely to call for a vote than all others. In *Rule 8*, consistent proselfs are more likely to call for a vote than all others.

Note that we are not arguing that the proselfs (prosocials) never have a conflict in *Rule 2* (*Rule 8*). For example, our discussion following Theoretical Prediction 1 (applied to the experimental environment in fn. 16) makes clear when proselfs might want to change *Rule 2*. Like TP1, *H1* is a comparative static prediction comparing the likelihood of rule change attempt between the types.

H2 (based on TP2).

- a) In *Rule 2*, consistent prosocials will first try to change the existing norm to higher contributions than five.¹⁷
- b) If their norm-change attempt does not succeed, consistent prosocials will call for a vote.

H3 (based on TP3).

The existing *Rule* is changed more often when communication is allowed.

¹⁷ Note that in *Rule 8*, individuals cannot try to change the norm to be below 5 because contributions of 8 points are enforced.

4. Results

4.1. Descriptives¹⁸

Social values

Our two measures of social value orientations allow us to assess whether social values are affected by interactions and behavioral outcomes in PGG1. Of the 215 participants, 174 (81%) were categorized equally at time 1 (SVO1) and time 2 (SVO2). For this reason, we base our analyses on the first measure, SVO1.

Our categorization reveals that no participant made competitive choices, leaving only prosocials and proselfs. 20 participants (9.3%) remain uncategorized because they made fewer than six consistent choices. As is customary when applying this SVO measure (Stouten, De Cremer, and Van Dijk 2005), these participants are not considered in the analyses. This leaves us with a total of 195 participants (112 men and 83 women, $M_{\text{age}} = 22.65$, $SD = 3.24$). Of these, 94 (48.2%) are *consistently* proself and 52 (26.7%) are *consistently* prosocial.

We created a variable *Conflict* between social values and the existing contribution rule, that indicates when a participant's value orientation is opposite to what the rule requires. Table 1 summarizes how this variable is created, and shows that 28 (23.5%) out of 119 participants in the *Rule 2* treatment have a value-rule conflict. These participants are consistently prosocial while the formal rule is 'proself'. In comparison, 34 (44.7%) out of 76 participants in the *Rule 8* treatment have a consistent proself value orientation while facing a 'prosocial' formal rule. We find no significant differences between the fractions of consistent proselfs (and prosocials) in *Rule 2* versus *Rule 8* (Fisher exact test, $p = .177$).

¹⁸ See online Appendix B for a discussion of the socio-demographics of our participants

(<https://www.eui.eu/Documents/DepartmentsCentres/SPS/Profiles/Gerxhani/Appendix-AandB.pdf>).

Table 1. Value-Rule conflict

Rule	Norm	Social value	Value-rule conflict	Conflict category
(Rule 2) Low level contribution (N = 119)	5	Consistently Prosocial	Yes (N = 28)	1
		Consistently Proself	No (N = 60)	3
		No Consistent Social Value	No (N = 31)	3
(Rule 8) High level contribution (N = 76)	5	Consistently Prosocial	No (N = 24)	3
		Consistently Proself	Yes (N = 34)	2
		No Consistent Social Value	No (N = 18)	3

Note: The final column assigns individuals into one of three categories: (1) prosocial individuals with a value-rule conflict because they have a consistent value (column 3) that conflicts with the *Rule 2* (column 1); (2) proself individuals with a value-rule conflict because they have a consistent value (column 3) that conflicts with the *Rule 8* (columns 1); and (3) participants without a value-rule conflict. Participants who cannot be classified as either prosocial or proself (N = 20) are not included.

Social norm

Because we make the social norm of contributing five points common knowledge, the aggregate group contributions revealed at the end of a round in PGG1 provides participants with a first indication of the extent of norm compliance in the team. Those who do not conform to the social norm might be signaling a desire for an alternative norm (Knight and Enslinger, 1998). We therefore consider individual contributions in the first round as a signal to the team members of behavior that deviates from the shared norm. We thus measure *norm-change attempts* as first-round individual contributions in excess of the norm of five points.¹⁹

The norm of contributing five points was correctly estimated by 30% of the pre-study participants.²⁰ An individual's belief about the prevalent norm is not affected by her own

¹⁹ In the *Rule 2* treatment we consider upward deviations from five points because of our focus on a rule change from a minimum contribution of two to eight points. In *Rule 8* our focus is on a rule change from eight to two points. As mentioned above, downward deviations from five are not possible due to strict rule enforcement.

²⁰ Online Appendix B shows the distribution of these estimates.

values; though prosocials report a slightly higher norm ($M = 6.58$, $SD = 2.01$) than proselfs ($M = 6.43$, $SD = 2.31$), the difference is far from statistical significance (two-tailed Mann-Whitney, $z = -0.912$, $p = .362$).

4.2. Institutional Change

Values

Our analysis begins with a test of H1, which predicts that consistent prosocials are most likely to attempt to change *Rule 2* while consistent proselfs are most likely to try to change *Rule 8*.

We start with the *Rule 2* treatment to investigate whether having a conflict affects the likelihood of calling for a vote to change *Rule 2* to *Rule 8*. Table 2 provides the results of a logistic regression of the call to vote on being consistently prosocial (that is, having a value-rule conflict), risk attitude, gender and age. The results reveal that a value-rule conflict is a strong and significant predictor of a rule-change attempt from *Rule 2* to *Rule 8* (second and third column). None of the controls has a significant effect on the call to vote.

Table 2. Rule-change attempts

	<i>Rule 2</i>	SE	<i>Rule 8</i>	SE
Conflict (consistently prosocial)	.976*	.463	.775	1.187
Risk	.200	.130	1.817**	.639
Age	-.055	.069	-.021	.198
Gender = Female	.541	.413	2.460	1.361
Constant	-.909	1.788	-17.142	7.051
N	119		79	
Nagelkerke R ²	8.6%		50.7%	

Notes: In Rule 2 (8) the consistently prosocial (proself) have a conflict.

*/** $p < 0.05/p < .01$ (two-tailed test).

A similar analysis for the *Rule 8* treatment shows no significant effect of a value-rule conflict on the attempts to change to *Rule 2* (forth andf fifth column of Table 2), that is, consistent

proselfs are not more likely to attempt change than all others. The regression also shows that attempting this rule change is positively correlated with one's willingness to take risks.

The results in table 2 show that the role of a value-rule conflict in individual attempts to change formal rules depends on the institutional environment. When the latter is based on a formal rule requiring a low level of contribution, we find that a value-rule conflict is the main driver of formal institutional change. Hypothesis 1 is thus partially supported.

Social norm

H2a predicts that consistent prosocials in *Rule 2* will first try to change the existing norm. A norm-change attempt refers to first-round contributions above the social norm of five points. We have 28 cases with a value-rule conflict and 91 without (cf. table 1). Of those with a conflict (the consistent prosocials), 17 (60.7%) attempted to change the norm, while 34 (37.4%) of those without a conflict attempted to do so.²¹ This difference in proportions is significant at the 5% level (Fisher's exact test, $p = .048$). Therefore, we conclude that individuals having a value-rule conflict do indeed attempt to change the social norm first, thus providing support for H2a.

At first sight, one might think that this result simply replicates the known effect that prosocials contribute more to a public good than proselfs (e.g., Offerman *et al.*, 1996, Van Dijk *et al.*, 2002). This is not obviously the case, however, because there are only two possible motivations for prosocials to do so, neither of which can explain what we observe in our data. On the one hand, prosocials might unconditionally contribute more than proselfs, e.g., for reasons of pure altruism. If this were the case, they would not only contribute in the

²¹ Recall that those without a conflict include prosocials who were not completely consistent in choosing the prosocial option in SVO1. It is possible that some of these individuals also experienced some conflict between their social value and the existing contribution rule.

first round of PGG, but in all rounds. This is not observed in our data. On the other hand, prosocials might start to contribute and continue to do so, conditional on others contributing. In our design, such ‘conditional cooperators’ (Keser and van Winden 2000) have no reason to contribute fully in the first round, however, because they have been informed that the social norm is to contribute five points. We conclude that our observation that prosocials contribute more than five points can be interpreted as signaling a desire to change the social norm.

Next, we investigate whether a failed norm-change attempt affects the likelihood of a subsequent attempt to change the existing formal rule (H2b). We consider a norm change to be successful if groups reach a mean contribution of eight points or higher in all ten rounds of PGG1.²² In *Rule 2*, attempts to change the norm were observed in 25 of the 26 groups. Only three groups succeeded in this attempt (12%). Of the 14 individuals involved in these three groups, only two called for a vote after PGG1 as opposed to 47 of the 105 participants in groups where the norm did not change.

H2b predicts that individuals with a conflict try to change the contribution rule directly via voting, if they previously tried to change the norm and failed. To test this conjecture, we only consider teams where the norm did not change. This reduces the number of observations to 105. Table 3 provides the results of a logit regression of the call to vote (*Rule 2* to *Rule 8*) on the same set of regressors as in table 2, and includes two dummy variables and their interaction. The first reflects participants with a conflict (that is, the consistently prosocial) and the second captures individuals who have attempted a (failed) norm change. The interaction term between these two dummies allows us to isolate the group that H2b is concerned with. Including this interaction term means that the first dummy

²² We chose this threshold because it implies that the minimum contribution level of eight points is reached without a formal rule. This is a rather high threshold, as one could argue that it may take more rounds than ten for norm-change attempts to succeed. Note that our choice makes it more difficult to find support for H2.

reflects individuals with a conflict who did not attempt a norm change, while the second concerns individuals without a conflict who attempted a (failed) norm change. For these two dummies, we do not expect an effect on initiating rule change. Indeed, the coefficients are not significantly different than zero. The hypothesis is concerned with individuals with a conflict who have attempted a norm change and failed. To test the effect on initiating rule change, we need to test whether the sum of the coefficients for individuals with conflict, for those who have attempted to change the norm and failed, and for the interaction term is different from zero. We find strong evidence that this is the case; the null of no effect is rejected ($\chi^2 = 8.56$; $p = .003$). This evidence that attempts to change the rule are preceded by attempts to change the norm provides strong support for H2b.

Table 3. Direct rule change attempts (*Rule 2 to Rule 8*)

	B	SE
(1) Conflict (consistently prosocial)	.386	.757
(2) (Failed) Norm-Change Attempt	.0742	.483
(3) Conflict x (Failed) Norm-Change Attempt	1.058	1.077
Risk	.239 ⁺	.145
Age	-.045	.074
Gender = Female	.468	.446
Constant	-1.483	1.956
(1)+(2)+(3)=0	$\chi^2 = 8.56$	$p = .003$
Nagelkerke R ²	16.4%	

Notes: $N = 105$; ⁺ $p < .10$ (two-tailed tests).

Communication

Finally, we expected a change of the contribution rule to be facilitated by communication (H3). Whether communication has this effect may depend on the institutional environment. For example, participants in *Rule 8* might feel hesitant to communicate because their wish to change it to *Rule 2* would signal a proself value (which they may not want to share in a ‘prosocial’ environment). We therefore considered the two rule treatments separately. To fully utilize the individual-level data that we have, we consider individual votes on whether

to change the rule (treating the cases where no one called to vote as a vote against the change). This approach is based on the idea that the likelihood of a vote in favor of change and the likelihood of an actual rule change are highly correlated.

In *Rule 2*, 79.7% of the (59) participants in the treatment without communication are in favor of a rule change, and 86.7% of the (60) participants in the treatment with communication are in favor. The difference is not significant at conventional levels (Fisher's exact test, $p = .337$). The percentages are, respectively, 7.9% and 2.6% in *Rule 8*, but again not significant (Fisher's exact test, $p = .615$). These results do not support H3.

5. Concluding discussion

We address agency-driven institutional change by experimentally investigating the micro-processes of individuals' social values, social norms and actions within the constraints and opportunities implied by existing formal rules. In particular, we consider two formal institutional environments that vary with respect to the extent of mandatory cooperation (i.e., the extent to which an existing rule leaves the decision to act cooperatively at the discretion of the individual). Doing so provides empirical evidence to the macro-micro-macro link in analytical sociology (Coleman, 1990). Our results show that formal institutional change is most likely to be initiated by individuals who experience a conflict between prosocial values and existing rules allowing for low cooperation.

Moreover, if such individuals with a value-rule conflict face a social norm they disapprove of, they start their attempts at institutional change by trying to change this social norm. They do so by signaling a desire to deviate from it. Note that attempting a norm change by deviating from it is not *a priori* obvious. This is because doing so may be costly due to possible social pressure. In our experiment, deviating from the social norm in a more

prosocial direction is costly because of the nature of the public goods game: while the income of other team members increases, the deviator's income decreases.

The results show that although individuals make many attempts to change the social norm, they often fail. When this occurs, individuals try to change the formal rule directly. In our experiment this is undertaken by calling for a vote to decide about a rule change. Note that a rule-change attempt after a failed norm-change attempt comes at an additional loss of income because of the tangible costs associated with it (e.g., in our experiment, these are the costs related to calling for a vote). Nevertheless, many individuals in our study not only deviate from the existing social norm in an attempt to change it, but also follow through with a rule-change attempt after the norm change has proven unsuccessful. The results also show that for the few cases where a norm-change attempt was successful, it was rarely followed by a subsequent rule-change attempt. This confirms theoretical arguments proposed in the literature that new social norms can become self-reinforcing and can render a change in formal rules unnecessary (Nee and Ingram, 1998; Ostrom, 1990). Taken together, these findings are in line with Hodgson's (2006) conclusion that while some institutions can be self-enforced, others need a more formalized sustainment.

We do not find evidence that communication facilitates an actual rule change, which is rather surprising because communication between team members has been shown to solve cooperation problems (Janssen *et al.*, 2014; Ostrom, 2005). In fact, we observe that a majority of those voting are in favor of a formal rule change towards more mandatory cooperation, irrespective of whether they were able to communicate beforehand. One possibility is that participants are familiar with the environment characterizing the social dilemma. A call to vote for a change in the less cooperation rule may be informative in itself without the need for more discussion. If this is the case, communication will have no effect. More research is needed to further investigate the role of communication in this setting.

There is an asymmetry in our results across the two formal institutional environments that we examine. Rule-change attempts are both more frequent and more successful when changing a rule from low to high cooperation than vice versa. Very few individuals with a conflict between their proself values and a (prosocial) high cooperation rule attempt to change the latter. This result may not be surprising if we consider the nature of the public good game in our experiment. A proself value implies acting in a way that maximizes the own (economic) well-being. In our study, the formal rule requiring high cooperation ensures not only an equal and prosocial distribution of wealth but also high earnings for all. A change to a less stringent rule might reduce the earnings even of a proself individual who would typically contribute less after the change. This could motivate her to not initiate a change and maintain the high cooperation rule. On the other hand, proself individuals will want to reduce the minimum contribution level if they expect to benefit from doing so. This will be the case if proselfs expect that sufficient other team members will maintain high contribution levels even after the required minimum has been reduced. In that case, free-riding on the prosocials' contributions would be a profitable alternative to being forced to contribute highly. The high cooperation rule, however, does not allow the proselfs to collect information about others' willingness to contribute under less restrictive contribution rules. In short, while prosocials in the low contribution requirement can contribute above the minimum to see how others respond, proselfs cannot contribute less when high contribution is mandatory. This asymmetry caused by the rules might explain the asymmetry in our results.

We have focused here on institutional change within teams, the simplest form of an organization. The laboratory environment allows us to draw causal conclusions on underlying mechanisms of institutional change, a task that would be more difficult under more complex organizational structures and virtually impossible with observational field data. This high internal validity leaves open the matter of the generalizability of our results (i.e., the external

validity). Many organizations outside of the laboratory are more involved than teams and do not allow for detailed analysis of the complex interactions between values, norms, and rules. Our study intends to provide a benchmark for further investigation under more complicated settings.

One line of future research may systematically introduce new features into our laboratory design to directly check the robustness of our results. For instance, while we believe the public goods game to be a well-suited tool for studying the agency of institutional change in social dilemma situations, institutional change might have different patterns in other interactive environments. Similarly, our experiment enables individuals to directly vote on formal institutional change. Further research should also aim to investigate institutional change in alternative decision-making environments. Another line of future research could involve moving from the laboratory to the field. This can be done by running field experiments along lines similar to our design, or by collecting observational data in the field (for example via surveys amongst members of organizations). The latter will not allow for conclusions about causality, but their high external validity would provide a powerful addition to the causal inferences we have been able to derive from our experiments.

Finally, by investigating the interplay between agency and institutional structure, our study was able to empirically show that the institutions-as-rules and the institutions-as-equilibria perspectives “become entwined” (Hodgson, 2006: 21). The main takeaway of our study may well be the importance of individuals, their (non)material motives, and group-level processes for understanding the evolution of rules in organizations.

Acknowledgements

Financial support from the NWO (ASPASIA-VIDI #015.007.048) is gratefully acknowledged. We would like to thank seminar participants at the 2014 ESA International

Meetings in Hawaii and the participants at the 2015 IMEBESS conference in Toulouse who contributed to earlier versions of this work with helpful comments.

References

- Bicchieri, C. (2006), *The Grammar of Society: The Nature and Dynamics of Social Norms*, New York: Cambridge University Press.
- Brousseau, E., P. Garrouste, and E. Raynaud (2011), 'Institutional changes: Alternative theories and consequences for institutional design', *Journal of Economic Behavior and Organization*, 79(1): 3-19.
- Chen, X., T. Gross, and U. Dieckmann (2013), 'Shared rewarding overcomes defection traps in generalized volunteer's dilemmas', *Journal of Theoretical Biology*, 335(21): 13-21.
- Coleman, J. (1990), *Foundations of social theory*, Cambridge: Harvard Univ. Press.
- Diekmann, A. (1985), 'Volunteer's dilemma', *Journal of Conflict Resolution*, 29: 605-610.
- DiMaggio, P. J. (1988), 'Interest and Agency in Institutional Theory', in L. G. Zucker (ed.), *Institutional Patterns and Culture*, Cambridge: Ballinger Publishing Company, pp. 3-22.
- Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011), 'Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences', *Journal of the European Economic Association*, 9(3): 522-550.
- Elster, J. (1989), 'Social Norms and Economic Theory' *The Journal of Economic Perspectives*, 3(4): 99-117.
- Fehr, E. and Simon Gächter (2000), 'Fairness and retaliation: The economics of reciprocity', *The Journal of Economic Perspectives*, 14(3): 159-181.
- Giddens, A. (1984), *The constitution of society*, Berkeley: University of California

Press.

Greif, A. and C. Kingston (2011), 'Institutions: Rules or Equilibria?', in N. Schofield and G.

Caballero (eds.), *Political Economy of Institutions, Democracy and Voting*,

Berlin/Heidelberg: Springer Berlin Heidelberg, pp. 13-43.

Hauser, O. P., D. G. Rand, A. Peysakhovich, and M. A. Nowak (2014), 'Cooperating with

the Future', *Nature*, 511(7508): 220–223.

Hitlin, S. and K. Pinkston (2013), 'Values, Attitudes, and Ideologies: Explicit and

Implicit Constructs Shaping Perception and Action', in J. DeLamater and A. Ward

(eds.), *Handbook of Social Psychology*, Dordrecht: Springer Netherlands, pp. 319–

339.

Hodgson, G. (2006), 'What are Institutions', *Journal of Economic Issues*, 40(1): 1-25.

Holt, C. A. and S. K. Laury (2002), 'Risk Aversion and Incentive Effects',

American Economic Review, 92:1644-1655.

Jackson, M. and D. R. Cox (2013), 'The Principles of Experimental Design and Their

Application in Sociology', *Annual Review of Sociology*, 39: 27–49.

Jacob, D. (2015), 'Every Vote Counts: Equality, Autonomy, and the Moral Value of

Democratic Decision-Making', *Res Publica*, 21(1): 61–75.

Janssen, M., R. Goldstone, F. Menczer, and E. Ostrom (2008), 'Effect of rule

choice in dynamic interactive spatial commons', *International Journal of the*

Commons, 2(2): 288-312.

Janssen, M., M. Tyson, A. Lee (2014), 'The effect of constrained communication

and limited information in governing a common resource', *International Journal of*

the Commons, 8(2): 617-635.

Keser, C. and F. van Winden (2000), 'Conditional Cooperation and Voluntary

Contributions to Public Goods', *The Scandinavian Journal of Economics*, 2(1): 23-39.

- Kingston, C. and G. Caballero (2009), 'Comparing Theories of Institutional Change', *Journal of Institutional Economics*, 5(2):151–180.
- Knight, J. and J. Ensminger (1998), 'Conflict over Changing Social Norms: Bargaining, Ideology, and Enforcement', in M. C. Brinton and V. Nee (eds.), *The New Institutionalism in Sociology*, Stanford: Stanford University Press, pp. 105–126..
- Krupka, E. and R. A. Weber (2013), 'Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?', *Journal of the European Economic Association*, 11(3): 495–525.
- Miller, L. M. (2008), 'Two notions of conventions: An experimental analysis', *Journal of Institutional Economics*, 4(3): 327–349.
- Morrison, E. W. (2006), 'Doing the Job Well: An Investigation of Pro-Social Rule Breaking', *Journal of Management*, 32(1): 5–28.
- Nee, V. and P. Ingram (1998), 'Embeddedness and Beyond: Institutions, Exchange and Social Structure', in M. C. Brinton and V. Nee (eds.), *The New Institutionalism in Sociology*, Stanford: Stanford University Press, pp. 19–45.
- North, D. C. (1990), *Institutions, Institutional Change and Economic Performance*, Cambridge, UK: Cambridge University Press.
- Offerman, T., J. Sonnemans, and A. Schram (1996), 'Value Orientations, Expectations and Voluntary Contributions in Public Goods', *The Economic Journal*, 106(437): 817–845.
- Olson, M. (1965), *The Logic of Collective Action: Public Goods and the Theory of Groups*, Cambridge, MA: Harvard University Press.
- Ostrom, E. (1990), *Governing the Commons: The Evolution of Institutions for Collective Action*, Cambridge, UK: University Press.

- Ostrom, E. (2005), *Understanding Institutional Diversity*, Princeton: Princeton University Press.
- Parsons, T. (1966), *Societies: Evolutionary and Comparative Perspectives*, Englewood Cliffs: Prentice Hall.
- Pierson, P. (2000), 'Increasing Returns, Path Dependence, and the Study of Politics', *The American Political Science Review*, 94(2): 251-267.
- Rand, D. G. and M. A. Nowak (2013), 'Human Cooperation', *Trends in Cognitive Sciences*, 17(8): 413–425.
- Reuben, E. and A. Riedl (2013), 'Enforcement of Contribution Norms in Public Good Games with Heterogeneous Populations', *Games and Economic Behavior*, 77(1):122–137.
- Sefton, M., R. Shupp, and J. Walker (2007), 'The effects of rewards and sanctions in the provision of public goods', *Economic Inquiry*, 45: 671-690.
- Sewell, G. (1998), 'The Discipline of Teams: The Control of Team-Based Industrial Work Through Electronic and Peer Surveillance', *Administrative Science Quarterly*, 43(2): 397–428.
- Stouten, J., D. De Cremer, and E. Van Dijk (2005), 'All Is Well That Ends Well, at Least for Proselfs: Emotional Reactions to Equality Violation as a Function of Social Value Orientation', *European Journal of Social Psychology*, 35(6): 767–783.
- Tao, L. and W. Au (2014), 'Values, Self and Other-Regarding Behavior in the Dictator Game', *Rationality and Society*, 26(1):46–72.
- Turner, J. C., P. J. Oakes, S. A. Haslam, and C. McGarty (1994), 'Self and Collective: Cognition and Social Context', *Personality and Social Psychology Bulletin*, 20(5): 454–463.
- Van Dijk, F., J. Sonnemans, and F. van Winden (2002), 'Social ties in a public good experiment', *Journal of Public Economics*, 85: 275–299.

- Van Lange, P., E. De Bruin, W. Otten, and J. Joireman (1997), 'Development of Prosocial, Individualistic, and Competitive Orientations: Theory and Preliminary Evidence', *Journal of Personality and Social Psychology*, 73(4): 733–746.
- Voigt, S. (2018), 'How to measure informal institutions', *Journal of Institutional Economics*, 14(1): 1-22.
- Yamagishi, T. (1986), 'The Provision of a Sanctioning System as a Public Good', *Journal of Personality and Social Psychology*, 51(1):110–116.
- Walker, J. M., R. Gardner, A. Herr, and E. Ostrom (2000), 'Collective Choice in the Commons: Experimental Results on Proposed Allocation Rules and Votes', *The Economic Journal*, 110(460): 212–234.
- Weesie, J. and A. Franzen (1998), 'Cost Sharing in a Volunteer's Dilemma', *Journal of Conflict Resolution*, 42(5): 600-618.

Appendix A. Instructions Main Experiment.

You are participating in an experiment. In this experiment, you can earn money. How much you earn depends on the decisions you make and on the decisions the other participants make in this experiment. In addition, you will receive a €7 show-up fee.

During the experiment you will gain points. Your earnings in points will be paid in Euros at the end of the experiment. This payment will be made confidentially to one participant at a time. The exchange rate used is 100 euro cents to 100 points.

It is important that you fully understand the instructions. For this reason, we ask you to read them carefully. In these instructions numerical examples are used. These serve only as an explanation; they have no special meaning with regard to the experiment itself.

Talking to other participants is not allowed during the experiment. If you have a question regarding any of the experimental tasks please raise your hand. A CREED employee will come by to assist you.

There will be many pages of instructions. You can move from one page to the next by clicking (with the mouse) on "next".

The experiment consists of two parts. We begin with the instructions for part 1.

Part 1 consists of three tasks.

During the experiment, we keep you informed about your progress, so that you know which task you are engaged in.

We now start with task 1. It involves answering a few questions.

In this task, there are no right or wrong answers.

Select the number that best describes how you feel right now. Do not reflect for too long, but give

your immediate reaction.

This is the end of the instructions for task 1 of part 1.

When you've read these instructions, click on "done." When everyone is ready, we will start with task 1.

[Emotions 1 Measure]

Please read the following descriptions and indicate how much they are applicable to you right now, meaning *at this very moment*.

I feel anger

Not at all	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very much
------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------

I feel shame

Not at all	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very much
------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------

I feel guilt

Not at all	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very much
------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------

I feel gratitude

Not at all	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very much
------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------

I feel regret

Not at all	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very much
------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------

[Risk Measure]

To conclude, we would like you to answer this question:

How willing are you to take risks, in general?

Not at all ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very much

Everyone has now completed task 1 of part 1.

We continue with the second task of part 1.

During this task you are randomly paired with another person whom we will refer to as the “Other”.

This “Other” person is someone you do not know and who you will not meet in the future.

During this task both you and the “Other” person will choose nine times between three options, which we refer to as A, B, and C.

Your own choices will generate points for both yourself and the “Other” person. Likewise, the “Other's” choices will generate points for him/her and for you. Every point has a value: the more points you receive, the better for you, and the more points the “Other” receives, the better for him/her. The choices of the “Other” do not affect your earnings; they affect the earnings of one of the other participants to whom (s)he is matched. However, you can also earn points because you are the “Other” for someone else in the lab.

An example to illustrate:

- You are matched with “Robin”. Your choices determine the earnings of yourself and Robin.
- Robin is matched with “Sam”. Robin’s choices determine the earnings of Robin and Sam.
- Finally, “Renee” is matched with you. Renee’s choices determine the earnings of Renee and yourself. So, the points you earn are determined by Renee’s and by your own choices.

Here is an example of how this task works:

	A	B	C
You get	500	500	550
The other gets	100	500	300

In this example, if you choose option A you would receive 500 points and the other would receive 100 points; if you chose B, you would receive 500 points and the other 500 points; and if you choose C, you would receive 550 points and the other 300 points.

You can thus see that your choices affect both the number of points you receive and the number of points the other receives.

In this same way you will choose nine times between A, B, and C options. One of these choices will be randomly selected to be paid out at the end of the experiment. This holds for each participant. You learn at the end of the experiment which of the choices is selected, and how much you have earned during this task.

Before you begin to make your choices, please keep in mind that there are no right or wrong answers. Choose the option that you prefer the most.

This is the end of the instructions for task 2 of part 1.

When you have read these instructions, click on "done." When everyone is ready, we will start with task 2.

[Triple dominance measure 1 (SVO1)]

For each of the nine situations below, choose A, B, or C, depending on which column you prefer the most:

A	B	C
---	---	---

(1) You get	480	540	480
The other gets	80	280	480
	O	O	O

	A	B	C
(2) You get	560	500	500
The other gets	300	500	100
	O	O	O

	A	B	C
(3) You get	520	520	580
The other gets	520	120	320
	O	O	O

	A	B	C
(4) You get	500	560	490
The other gets	100	300	490
	O	O	O

	A	B	C
(5) You get	560	500	490
The other gets	300	500	90
	O	O	O

	A	B	C
(6) You get	500	500	570
The other gets	500	100	300
	O	O	O

	A	B	C
(7) You get	510	560	510
The other gets	510	300	110
	O	O	O

	A	B	C
(8) You get	550	500	500
The other gets	300	100	500
	O	O	O

	A	B	C
(9) You get	480	490	540
The other gets	100	490	300
	O	O	O

Everyone has completed task 2 of part 1.

We continue with the third and last task of part 1.

A few weeks ago, other people, who are registered at the creedexperiment.nl pool, participated in a study. Firstly we will describe to you the situation that was presented to them, as well as the choice they were asked to make. Then, we will ask you to predict what these people chose.

Please read the instructions carefully and try to empathize with these participants.

What follows are the instructions that were presented to these participants. These instructions consist of several pages. You can start by clicking on "next page".

Imagine the following situation:

You are part of a team of five people. Each team member receives an income of 10 points. You have to decide how many of these 10 points you want to contribute to a team project and how many you want to keep for yourself. Each team member must make this same decision.

The contributions are added up and multiplied by two by the organizers of this experiment.

Thereafter, the pot is divided equally among the five team members. For every euro contributed you get paid $2/5 = 0.4$ euro, regardless of who made the contribution.

Your contribution to the team project thus leads to an income increase for the other team members.

On the other hand, you also earn income through the contribution of the other team members to the team project.

Your income (in points) is therefore:

$(10 \text{ points} - \text{your contribution to the team project}) + 0.4 * (\text{total contribution to the team project})$.

The income of each team member is calculated in the same way, so that each team member receives the same income from the team project.

Below three examples are given, where each team member has an income of 10 points per round:

1) You contribute 8 points to the team project. Every other team member also contributes 8 points. In total, 40 points are contributed.

Your income is $(10-8) + 0.4 * 40 = 18$ points. Also the income of each individual team member is 18 points.

2) You contribute 2 points to the team project. Every other team member also contributes 2 points. In total, 10 points are contributed.

Your income is $(10-2) + 0.4 * 10 = 12$ points. Also the income of each individual team member is 12

points.

3) You contribute 5 points to the team project. Two other team members contribute 2 points each whereas the last two contribute 8 points each. In total, 25 points are contributed.

Your income is $(10-5) + 0.4 * 25 = 15$ points. The income of each team member who contributed 2 points is 18 points. The income of the other two team members who contributed 8 points is 12 points each.

To sum up: each participant may contribute 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 points to the team project.

We asked the participants to indicate which contribution to the team project they found “socially appropriate” and “in accordance with moral and correct social behaviour”.

With “socially appropriate” behaviour we mean behaviour that most people consider as the correct, proper and ethical behaviour.

If a contribution decision of a participant (a number between 0 and 10) matched the modus of other decisions, this participant received a bonus.

Now we want to know from you what you think the modal response was of this group of participants. The modal response is the most frequently chosen response by the participants. In this case, it is the modal response given by participants a few weeks ago.

If your answer matches the modal response of this group of participants you will receive a bonus of 500 points.

[Norm measure]

Indicate which contribution to the team project (a number between 0 and 10) was most frequently mentioned as “socially appropriate”.

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
0	1	2	3	4	5	6	7	8	9	10

Before we continue with the instructions of part 2, we will give you feedback on the previous task.

The modal response (the answer that most participants gave) of the participants in the study a few weeks ago was 5.

This means that the contribution that is most frequently mentioned as “socially appropriate” is 5.

This brings us to part 2 of the experiment.

Part 2 consists of 4 tasks.

The first and the last tasks consist of 10 rounds each. The second and the third tasks consist of a short assignment.

During the first and the last tasks you will be grouped with four other participants. Together you form a team of five members.

The composition of this team will not change: you will remain with the same five members in the same team for all 20 rounds of these two tasks. You will remain anonymous during and after the tasks. Team members are assigned a number.

We now begin with the first task. Click on "next" to continue.

The first task consists of 10 rounds. In each one of these rounds, everyone in your team should decide about contributing to a team project. This occurs in exactly the same way as described in the instructions for the people who participated a few weeks ago.

To summarize, this is what will happen:

At the beginning of each round, each team member will receive 10 points. Every team member must decide how many of these points to contribute to the team project and how many points to keep for themselves.

Your income in one round:

$(10 \text{ points} - \text{your contribution to the team project}) + 0.4 * (\text{total contribution to the team project})$.

Each team member receives the same income from the team project, irrespective of what (s)he has contributed.

At the end of each round you will be informed as to the total contribution of all team members to the team project.

Click on "next" to continue

In the left hand corner of the screen you will soon see the round number and your total earnings of previous rounds.

To indicate how many points you want to contribute to the team project, enter the number (typically a number between 0 and 10) in the input field at the top center of the screen and confirm it.

By deciding on how many points you want to contribute to the team project you automatically decide

on how many points you keep for yourself. The number of points that you can contribute to the team project is 10.

Once you confirm your decision, you cannot change it.

After all team members have made their decision, the screen for the next round will appear.

At the center of the screen you will see your contribution and the overall team contribution of the previous round.

In addition, you will be able to see your earnings from the previous round (10 points - your contribution to the team project) + $0.4 * (\text{total contribution to the team project})$.

Finally, a special rule will be applied today:

In each of the 10 rounds, each team member must contribute at least 2 (8) points. You will see that it is impossible to enter a lower number than 2 (8).

If you have read and understood the instructions, click on "done" so we can start with the team project.

[PGG1]

This brings us to the end of the first task of part 2.

We now start with the second task of part 2.

This task again consists of answering a few questions.

There are no right or wrong answers.

Always select the number that best describes how you feel right now. Do not reflect too long but give your immediate reaction.

[Emotions 2]

Please read the following descriptions and indicate how much they are applicable to you right now, meaning *at this very moment*.

I feel anger

Not at all ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very much

I feel shame

Not at all ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very much

I feel guilt

Not at all ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very much

I feel gratitude

Not at all ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very much

I feel regret

Not at all ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very much

For the third task, you are re-paired with another person whom we call the “Other”. This is not the same person with whom you were paired in part 1. It is also unlikely that (s)he will be a member of your team of five. This “Other” person is someone you do not know and who you will not meet in the future.

As in part 1, both you and the “Other” person will choose nine times between three options: A, B, or C.

Your own choices will generate points for both yourself and the “Other” person. You also earn points because you are the “Other” for someone else in the lab.

To clarify, we repeat the example from part 1:

- You are matched with “Robin”. Your choices determine the earnings of yourself and Robin.
- Robin is matched with ”Sam”. Robin’s choices determine the earnings of Robin and Sam.
- Finally, “Renee” is matched with you. Renee’s choices determine the earnings of Renee and yourself. So, the points you earn are determined by Renee’s and by your own choices.

Before you begin to make your choices, please keep in mind that there are no right or wrong answers.

Choose the option that you prefer the most.

[Triple dominance measure 2 (SVO2)]

For each of the nine situations below, choose A, B, or C, depending on which column you prefer the most:

	A	B	C
(1) You get	480	540	480
The other gets 80	280	480	

O O O

	A	B	C
(2) You get	560	500	500
The other gets	300	500	100

O O O

A B C

(3) You get	520	520	580
The other gets	520	120	320
	O	O	O

	A	B	C
(4) You get	500	560	490
The other gets	100	300	490
	O	O	O

	A	B	C
(5) You get	560	500	490
The other gets	300	500	90
	O	O	O

	A	B	C
(6) You get	500	500	570
The other gets	500	100	300
	O	O	O

	A	B	C
(7) You get	510	560	510
The other gets	510	300	110
	O	O	O

	A	B	C
(8) You get	550	500	500
The other gets	300	100	500
	O	O	O

	A	B	C
(9) You get	480	490	540
The other gets	100	490	300
	O	O	O

This brings us to the final task of part 2.

During this final task you will be grouped again with four other participants. Together you will thus form the same team of five members as in the first task of part 2.

This final task of part 2 consists again of a team project of 10 rounds, similar to the first task of part 2.

Please remember that the rule which applied in the first 10 rounds of the team project was that team members had to contribute a minimum of 2 (8) points in each round.

We will now allow you to propose to your team members to vote for a change of this rule into a new rule, which ensures that each team member has to contribute a minimum of 8 (2) points in each round. Note that voting will take place only if at least one of the team members proposes to vote for a change of the rule. The one who proposes to vote bears a cost of 10 points. If several members make a proposal to vote, each one has a cost of 10 points.

If no one proposes to vote for a change of the rule, the same rule will apply as before. That means that each team member has to contribute a minimum of 2 (8) points in each round.

If at least one team member proposes to vote for a change of the rule, voting will take place:

- If no majority (i.e. two or fewer team members) supports the proposal, the same rule will apply as before, meaning that each team member contributes a minimum of 2 (8) points in each round.

- If a majority (i.e. three or more team members) supports the proposal, the rule will change and each team member must henceforth contribute at least 8 (2) points per round.

Participants in the Communication treatment read in this case the following instructions:

If at least one team member proposes to vote for a change of the rule, there is first the opportunity to have a discussion within the team via a chat box for 90 seconds. After this, voting will take place:

- If no majority (i.e. two or fewer team members) supports the proposal, the same rule will apply as before, meaning that each team member contributes a minimum of 2 (8) points in each round.

- If a majority (i.e. three or more team members) supports the proposal, the rule will change and each team member must henceforth contribute at least 8 (2) points per round.

Click on "ready" to start with the project team.

[Call to Vote Measure]

Do you want to propose to your team members to vote for a change of the rule?

☐ Yes

☐ No

[Anticipated regret Measure]

To which extent are the following statements applicable to your decision to vote or not for a change of the rule?

I did not propose to vote for a change of the rule because I did not want to feel bad if the result of the
vote was not as I want it to be.

Not at all ☐ ☐ ☐ ☐ Very much

I proposed to vote for a change of the rule because I did not want to regret not trying if the result of the vote were to turn out in my favor

Not at all ☐ ☐ ☐ ☐ Very much

[Vote Measure]

One or more of the team members has proposed to vote for a change of the rule.

Therefore you will now vote in favor or against the suggested rule change.

Participants in the Communication treatment read in this case the following instructions:

Before you vote, you have the opportunity to have a discussion with each other for 90 seconds.

Below a chat box will appear in which you can communicate for 1.5 minutes.

During the communication you cannot share your identity. Also you are requested to converse in a nice and polite manner.

Please indicate which of these rules you vote for.

- ☐ Every team member has to contribute 8 points each round
- ☐ Every team member has to contribute 2 points each round.

Remember that the majority (three or more members) determines the outcome.

We now begin with the last 10 rounds of the team project.

During these 10 rounds the following rule applies:

Each team member must contribute at least 2 (8) points per round. *(This depends on the result of the Vote)*

[PGG2] (*same as PGG1*)

Finally, we would like you to answer a few questions:

What is your age? _____

What is your gender? _____

Year of starting your study _____

Field of study _____

Do you have a paid job for more than 2 days a week? _____

During this experiment you had the opportunity to propose a vote for a rule change. Why did you or did you not choose to do so?

“I proposed to vote because.....”

“I did not propose to vote because.....”

Your answers will be treated confidentially.

This is the end of the questionnaire. Please remain seated and do not communicate until the other participants have finished this questionnaire.

We would like to thank you for your cooperation.

Appendix B: Additional Information

Socio-demographics

We checked all our variables of interest for differences across the socio-demographics of our participants. The only significant differences we found are as follows. First, men ($M = 6.71$, $SD = 1.51$) generally reported more willingness to take risks than women ($M = 5.94$, $SD = 1.70$), $t = 3.358$, $p = .001$). Second, age appears to correlate with mean contributions in PGG1 (ANOVA; $F(3, 193) = 4.040$, $p = .008$) and participants in a group of 20- to 23-year-olds more often show prosocial value orientations than in other age categories ($\chi^2(6, N = 194) = 7.985$, $p = .046$). To correct for these differences, we include gender and age as controls in our analyses.²³

Norm estimates

Figure B1 shows the distribution of estimates in the main experiment of the modal response in the pre-study. 34% of the participants correctly estimated the modal pre-study response to be five.

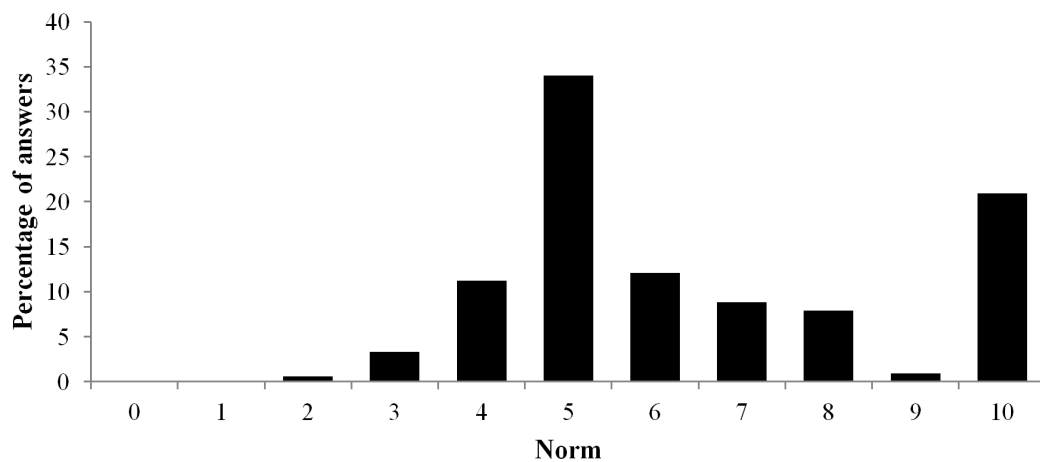


Figure B1. Norm measure.

Note: Bars show the distribution of answers (in percentages) in the main experiment to the question “Which contribution to the team project do you think the participants of our earlier experiment find most socially appropriate?”.

²³ Our results are not qualitatively affected if we exclude the corrections for age and gender.